

Application of reinforcement learning in mechatronic systems

Abstract: 本文目的是將實體機電系統簡化後導入虛擬環境並證明簡化後訓練的強化學習能算法能在模擬環境中應用。將實體冰球機的機電系統的雙自由度簡化成一個自由度導入 CoppeliaSim 模擬環境透過 Remote API 控制環境中的冰球機移動，OpenCV 來處理影像提供強化學習訓練和訓練後實際控制的輸入，強化學習訓練利用 OpenAI Gym 的 Pong Game 測試適合的訓練參數，再將算法套用到 CoppeliaSim 的場景中進行訓練。該研究為了證實相同的訓練算法能套用到不同真實程度的環境中進行訓練（，相同訓練參數套用到模擬環境中進行訓練，依舊能訓練出像樣的對打系統。）

Keywords: 類神經網路、強化學習、CoppeliaSim、OpenAI Gym

Chapter 1

介紹

近年來硬體技術、軟體、自動求導等技術快速發展起來，再次帶起機器學習的發展，促使機器學習與各領域結合的應用越來越廣泛，在機電系統採用強化學習是為了讓機電系統的控制達到最佳化。本研究利用強化學習優化冰球機的對打系統，並測試相同演算法運用在 2D 與 3D 模擬環境可行性。

本研究分成兩大部分，第一部分簡化冰球機並運用 OpenAI Gym 的 Atari Pong-v0 測試較合適的訓練參數，在 2D 環境中進行強化學習的對打訓練。第二部分將實體系統簡化後導入 CoppeliaSim 模擬環境導入測試完成的訓練參數進行虛擬訓練，進行算法導入到實體機前的測試。

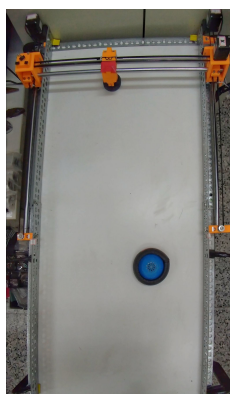


Figure 1.1: 實體的冰球機



Figure 1.2: 虛擬環境簡化後的冰球機



Figure 1.3: Gym 的 Pong game

Chapter 2

Literature Review

Atari Video Computer System 是 Atari Inc. 在 1977 年 9 月 11 日發行的一款家用遊戲機，直到 1982 年更名為 Atari 2600，Pong 是第一個主要發行的遊戲。OpenAI 是一個人工智慧研究實驗室，研究著重於強化學習，Gym 具有多種不同的環境，並以標準化定義環境，其中包括 Atari 2600 的遊戲、Classic control、Box2D 等環境供機器學習訓練使用。

冰球機與 Pong 都是透過玩家的眼睛觀察球的移動並透過操控擊錘進行攻擊與防守，並依照對手的球路來採取不同的策略。強化學習的訓練是透過影像處理，將擊錘及球的位置過濾出來，比較兩幀間擊錘與球的移動，透過強化學習判斷擊錘的移動，進行訓練。

Chapter 3

設計流程

3.1 簡化模型

將冰球機簡化成一個自由度，簡化後與 Pong 的操控相同，並分別透過 2D 遊戲 Pong-v0 與 3D 模擬環境 CoppilaSim 供強化學習訓練。

3.2 機器學習訓練

透過影像處理，過濾出擊錘與球，並將兩幀影像進行比較，透過調整環境參數與決策參數的權重，環境參數使用 ReLU Function 進行優化，決策參數則是透過 Log probability 和 Softmax 進行運算並在每局結束後透過 back-propagation 修正與更新偏差，藉由不斷的學習持續進步。

3.3 模擬環境訓練

在 CoppilaSim 中，抓取場景中的攝影機，